

CLAIMS

What is claimed is:

1. A speech data mining system for use in generating a rich transcription having utility in call center management, comprising:

a speech differentiation module differentiating between speech of at least two interacting speakers;

a speech recognition module improving automatic recognition of speech of a second speaker based on interaction of the second speaker with a first speaker employed as a reference speaker; and

a transcript generation module generating a rich transcript based at least in part on recognized speech of the second speaker.

2. The system of claim 1, wherein said speech differentiation module is adapted to receive speech input from the first speaker on a first channel, to receive speech input from the second speaker on a second channel, and to differentiate between the first speaker and the second speaker by identifying speech of the first speaker with speech received on the first channel, and identifying speech of the second speaker with speech received on the second channel.

3. The system of claim 2, wherein said speech recognition module is adapted to employ the first speaker as the reference speaker based on quality of the first channel being higher than quality of the second channel.

4. The system of claim 1, wherein said speech recognition module is adapted to employ the first speaker as the reference speaker based on availability of a speech model adapted to the first speaker.

5. The system of claim 1, wherein speech differentiation module is adapted to at least one of:

use a speech biometric trained on speech of the first speaker to distinguish between speech of the first speaker and speech of another speaker; and

use a speech biometric trained on speech of the second speaker to distinguish between speech of the second speaker and speech of another speaker.

6. The system of claim 1, wherein said speech recognition module is adapted to identify a topic with respect to which the speakers are interacting, and to employ a focused language model to assist in speech recognition based on the topic.

7. The system of claim 6, wherein said speech recognition module is adapted to receive an explicit topic selection from one of the speakers.

8. The system of claim 7, wherein said speech recognition module is adapted to prompt a speaker corresponding to a call center customer to explicitly select one of a plurality of predetermined topics by pressing a corresponding button of a telephone keypad.

9. The system of claim 7, wherein said speech recognition module is adapted to identify a predetermined topic associated with an electronic form selected by call center personnel.

10. The system of claim 6, wherein said speech recognition module is adapted to extract at least one keyword from a speech recognition result of at least one of the interacting speakers, and to identify a predetermined topic based on the keyword.

11. The system of claim 1, wherein said speech recognition module is adapted to extract context from a speech recognition result of the first speaker, and to employ the context extracted from the speech recognition result of the first speaker as context in a language model utilized to assist in recognizing speech of the second speaker.

12. The system of claim 1, wherein said speech recognition module is adapted to extract at least one keyword from a speech recognition result of the first speaker, and to supplement a constraint list used in recognizing speech of the second speaker based on the keyword extracted from the speech recognition result of the first speaker.

13. The system of claim 1, wherein said speech recognition module is adapted to extract at least one keyword from a speech recognition result of the first speaker, and to rescore recognition candidates generated during recognition of speech of the second speaker based on the keyword extracted from the speech recognition result of the first speaker.

14. The system of claim 1, wherein said speech recognition module is adapted to detect interruption of speech of one speaker by speech of another speaker, and to employ the interruption as context in a language model utilized to assist in recognizing speech of the second speaker.

15. The system of claim 1, wherein said speech recognition module is adapted to detect an interruption of speech of one speaker by speech of another speaker, and to record an instance of the interruption as mined speech data.

16. The system of claim 1, wherein said speech recognition module is adapted to extract at least one keyword from a speech recognition result of at least one of the interacting speakers, to identify a frustration phrase associated with the keyword, and to record an instance of the frustration phrase as mined speech data.

17. The system of claim 1, wherein said speech recognition module is adapted to extract at least one keyword from a speech recognition result of at least one of the interacting speakers, to identify a polity expression associated with the keyword, and to record an instance of the polity expression as mined speech data.

18. The system of claim 1, wherein said speech recognition module is adapted to extract at least one keyword from a speech recognition result of at least one of the interacting speakers, to identify a context corresponding to at least one of a topic, complaint, solution, and resolution associated with the keyword, and to record an instance of the context as mined speech data.

19. The system of claim 1, wherein said speech recognition module is adapted to identify a number of interaction turns based on a shift in interaction from speaker to speaker, and to record the number of turns as mined speech data.

20. The system of claim 1, comprising a quality management subsystem employing mined speech data as feedback to at least one of a call center quality management process and a consumptible quality management process.

21. A speech data mining method for use in generating a rich transcription having utility in call center management, comprising:

- differentiating between speech of at least two interacting speakers;
- improving automatic recognition of speech of a second speaker based on interaction of the second speaker with a first speaker employed as a reference speaker; and

- generating a rich transcript based at least in part on recognized speech of the second speaker.

22. The method of claim 21, wherein said step of differentiating between speech of at least two interacting speakers includes:

- receiving speech input from the first speaker on a first channel;
- receiving speech input from the second speaker on a second channel; and

- differentiating between speech of the first speaker and speech of the second speaker by identifying speech of the first speaker with speech received on the first channel, and identifying speech of the second speaker with speech received on the second channel.

23. The method of claim 22, comprising employing the first speaker as a reference speaker based on quality of the first channel being higher than quality of the second channel.

24. The method of claim 21, comprising employing the first speaker as a reference speaker based on availability of a speech model adapted to the first speaker.

25. The method of claim 21, wherein said step of differentiating between speech of at least two interacting speakers includes at least one of:

using a speech biometric trained on speech of the first speaker to distinguish between speech of the first speaker and speech of another speaker; and

using a speech biometric trained on speech of the second speaker to distinguish between speech of the second speaker and speech of another speaker.

26. The method of claim 21, wherein said step of improving automatic recognition includes:

identifying a topic with respect to which the speakers are interacting; and

employing a focused language model to assist in speech recognition based on the topic.

27. The method of claim 26, wherein the step of identifying a topic includes receiving an explicit topic selection from one of the speakers.

28. The method of claim 27, wherein said step of receiving an explicit topic selection includes prompting a speaker corresponding to a call center customer to explicitly select one of a plurality of predetermined topics by pressing a corresponding button of a telephone keypad.

29. The method of claim 27, wherein said step of receiving an explicit topic selection corresponds to identifying a predetermined topic associated with an electronic form selected by call center personnel.

30. The method of claim 26, wherein said identifying a topic includes:
extracting at least one keyword from a speech recognition result of
at least one of the interacting speakers; and
identifying a predetermined topic based on the keyword.

31. The method of claim 21, wherein said step of improving automatic recognition includes:

extracting context from a speech recognition result of the first speaker; and

employing the context extracted from the speech recognition result of the first speaker as context in a language model utilized to assist in recognizing speech of the second speaker.

32. The method of claim 21, wherein said step of improving automatic recognition includes:

extracting at least one keyword from a speech recognition result of the first speaker; and

supplementing a constraint list used in recognizing speech of the second speaker based on the keyword extracted from the speech recognition result of the first speaker.

33. The method of claim 21, wherein said step of improving automatic recognition includes:

extracting at least one keyword from a speech recognition result of the first speaker; and

rescoring recognition candidates generated by recognition of speech of the second speaker based on the keyword extracted from the speech recognition result of the first speaker.

34. The method of claim 21, comprising detecting an interruption of speech of one speaker by speech of another speaker, wherein said step of improving automatic recognition includes employing the interruption as context in a language model utilized to assist in recognizing speech of the second speaker.

35. The method of claim 21, comprising:
detecting an interruption of speech of one speaker by speech of another speaker; and
recording an instance of the interruption as mined speech data.

36. The method of claim 21, comprising:
extracting at least one keyword from a speech recognition result of at least one of the interacting speakers;
identifying a frustration phrase associated with the keyword; and
recording an instance of the frustration phrase as mined speech data.

37. The method of claim 21, comprising:
extracting at least one keyword from a speech recognition result of at least one of the interacting speakers;
identifying a polity expression associated with the keyword; and
recording an instance of the polity expression as mined speech data.

38. The method of claim 21, comprising:
- extracting at least one keyword from a speech recognition result of at least one of the interacting speakers;
 - identifying a context corresponding to at least one of a topic, complaint, solution, and resolution associated with the keyword; and
 - recording an instance of the context as mined speech data.
39. The method of claim 21, comprising:
- identifying a number of interaction turns based on a shift in interaction from speaker to speaker; and
 - recording the number of turns as mined speech data.
40. The method of claim 21, comprising employing the mined speech data as feedback to at least one of a call center quality management process and a consumptible quality management process.